

ABSTRACT

Systems and methods for logging and recovering updates to data structures in the event of failure of an information management system are provided. In exemplary implementations, methods for implementing an efficient redo log for a data structure that is concurrently accessed by multiple clients is provided. The data structure is implemented in two layers: the data structure algorithm layer which sits atop an allocator that provides distributed, persistent, and replicated storage allocation. Both the B-link tree algorithm layer and the allocator use the service of the logging mechanism to implement fault-tolerance and atomicity guarantees. The present invention uses a single log and allows periodic truncation of that log for space efficiency.